学校编码：10384　　分类号_____密级__

学号：21720071154155　　UDC___

厦门大学

硕 士 学 位 论 文

铜绿微囊藻 *PCC7806* 蛋白质的肽指纹图谱分析

The Use of MALDI-TOF Peptide Mass Fingerprinting to Study Proteins from Cyanobacterium *Microcystis aeruginosa PCC7806*

KAGIMBURA Ervine

指导教师姓名：章 军 教 授
专 业 名 称：生物化学与分子生物学

论文提交日期：2009 年 4 月
论文答辩时间：2009 年 5 月

学位授予日期：2009 年 月

答辩委员会主席：_____

评 人 阅：_____

2009 年 月

厦门大学学位论文原创性声明

兹呈交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文而产生的权利和责任。

声明人（签名）：

2009 年　　月

DECLARATION

I here declare that apart from the sources cited, this dissertation is my own work under the supervision of my supervisor Professor Zhang Jun. The materials of the dissertation have not been presented and it will not be presented to any other university other than this university for similar or any other degree.

Signature

Candidate: KAGIMBURA Ervine

厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

1、保密（ ），在　　年解密后适用本授权书。

2、不保密（ ）

（请在以上相应括号内打“√”）

作者签名：　　日期：2009 年 5 月　日

导师签名：　　日期：　年　月　日

COPYRIGHT

Signature

KAGIMBURA Ervine

目录表

Table of contents

List of tables

List of figures

摘要(Abstract in Chinese)

通过蛋白质组学，特别是肽质量指纹图谱的方法,我们鉴定出铜绿微囊藻 *PCC7806* 的蛋白。在研究中我们采用 2-DE 和 PMF 联用的策略。

铜绿微囊藻 *PCC7806* 培养于正常的 BG-11 培养基，通过超声破碎结合三氯乙酸/丙酮沉淀的方法提取藻全蛋白。双向电泳第一向等电聚焦采用 18cm IPG 胶条（pH 4-7），SDS–PAGE 为第二向。银染过的蛋白挖点，然后用胰酶消化。提取胰酶消化后的多肽片段与基质溶液混和，点样，装入 MALDI-TOF 质谱仪（BRUKER，ReflexTM III）作进一步分析。每个蛋白点获得的峰图经过了分析，以查明和消除污染峰和自切峰。接下来的步骤称为标峰，我们获得了 168 个蛋白点的质谱峰，将这些质谱峰提交至 NCBI/MASCOT 数据库以查找相匹配的蛋白质。通过蛋白图上的实验参数（pI 和分子量）与在线数据库搜索到的理论参数进行对比，我们实现了这些蛋白的搜索、鉴定。我们已经取得了以下成果：鉴定了 187 个蛋白质，并根据这些蛋白的来源物种和它们与我们研究的物种（铜绿微囊藻 *PCC7806*）的关系将它们分成 3 组。168 个蛋白点的质谱结果中,由于有 18 个点的匹配蛋白各包含 2 个,所以共得到 187 个蛋白。在这 187 个蛋白质鉴定结果中，有 14 个（7.5%）蛋白来源于铜绿微囊藻 *PCC7806*；有 98 个 （52.4%）的蛋白质来自属于蓝藻铜绿微囊藻 *PCC7806* 所属的蓝藻的进化树的不同分支； 有 75 个 （40.1％）的蛋白质，其分类与我们研究的铜绿微囊藻并不密切相关，但考虑其他的选择标准，如 pI 和分子量而被选择。与后两组数据相比,第一组中匹配蛋白的数量仍然较低,这说明还需要对铜绿微囊藻进行其他方面的研究.

关键词: 铜绿微囊藻 7806; MALDI-TOF; 肽指纹图谱; 双向电泳.

Abstract

We have identified proteins from *Microcystis aeruginosa PCC7806* by proteomic methods, especially by Peptide Mass Fingerprinting. The combined analysis by 2-DE and peptide mass fingerprinting strategies were used during this research.

Protein samples were extracted, by sonication and TCA/Acetone precipitation, from *Microcystis aeruginosa PCC7806* grown on normal BG-11 culture medium. 2-D electrophoresis was performed using IPG drystrips of pH 4-7 in the first dimension isoelectric focussing and SDS-PAGE in the second dimension. Proteins were stained by Coomassie or Silver staining prior to tryptic digestion. Trypsin porcine was used to digest our protein samples. Extracted peptides mixed with matrix solution were loaded on the plate and then installed on MALDI-TOF MS machine (BRUKER, ReflexTM III MALDI-TOF) for further analysis. Spectra obtained from different spots were analyzed in order to identify and remove contaminant picks. After this step known as calibration, pick reports of 168 protein spots were submitted to NCBI/ MASCOT database for identifying matching proteins. The on line searching has been realized by comparing experimental parameters (pI and molecular weight) estimated from the gel picture to theoretical parameters found from on line database.

From 168 spots, 187 proteins were identified since there are 18 spots containing 2 proteins for each spot. These proteins were grouped into three categories according to the relationship between their origin's taxonomy and the species of our case study *Microcystis aeruginosa PCC7806.* Of 187 identified proteins, 14 (7.5%) match to the organism of the case study; 98 (52.4%) proteins originate from other kind of cyanobacteria and 75 (40.1%) proteins are from the species which are not cyanobacteria but chosen by considering other criteria such as pI and molecular weight. The percentage represented in the first group show that the number of matching proteins is still low comparing with the last two groups; this means that other researches still need to be carried out on *Microcystis aeruginosa PCC7806*.

Keywords: *Microcystis aeruginosa PCC7806*; MALDI-TOF; Peptide Mass Fingerprinting; 2- D electophoresis.

List of abbreviations

1-D: one-dimensional electrophoresis

2-D: two-dimensional electrophoresis

ACN: acetonitrile

Acr: acrylamide

APS: Ammonium persulfate

ATCC: American Type culture Collection

BPB: bromophenol blue

CA: carrier ampholite

CCB: coomassie brilliant blue

CHAPS: 3-[(3-cholamidopropyl) dimethylaminol]-1-propanesulfonate

CID: collision induced dissociation

COGs: Clusters of Orthologous Groups

DTT: dithiothreitol

ESI: electrospray ionization

ESI-Q-IT-MS: electrospray ionization-quadrupole ion trap-MS

EST: Expressed sequence tags

EDTA: Etylenediaminetetraacetique acid

g: gram

Gly: glycine

HCCA/CHCA: α-cyano-4-hydroxycinnamic acid

IAA: iodoacetamide

IEF: Isoelectric focusing

IPG: Immobilized pH gradient

kDa: kiloDaltons

LC: Liquid chromatography

MALDI- TOF: Matrix Assisted Laser Desorption/Ionization- Time of Flight

MALDI-ToF/ToF: MALDI tandem-time-of-flight

Mb: mega base

min: Minute

Mowse: Moleculae weight search

MS: mass spectrometry

MW: Molecular weight

m/z: mass to charge ratio

NCBI: National Centre for Biotechnology Information

ng: nano gram

PAGE: polyacrylamide gel electrophoresis

PBS: Phosphate-Buffered Saline

PCC: Pasteur culture collection of cyanobacterial

pI: isoelectric point

PMF: peptide mass fingerprinting

PMSF: phenylmethanesulphonylfluoride

ppm: part per million

r/min: revolution per minute

QToF: quadrupole time-of-flight

SDS: sodium dodocyl sulfate

TAE: Tris-acetate-EDTA

TCA: trichloroacetic acid

TFA: trifluoroacetic acid

TEMED: N,N,N',N'-tetramethyl-ethylenediamine

Tris: tris (hydroxymethyl) aminomethane

WHO: World Health Organization

Chapter 1: Introduction

Protein identification from different organisms at different taxonomic levels has both scientific and practical value. Strain, species and genus specific proteins can provide insight into the criteria that define an organism and its relationships with close relatives. Such proteins can also serve as taxon specific diagnostic targets. Proteins are the main actionners in cells and in an entire organism [1]. Without proteins the most basic functions of life could not be carried out.

The function of proteins as enzymes is one of their main known functions. Enzymes are catalysts; they initiate a reaction between themselves and another protein, working on the molecule to change it in some ways. Different functions of proteins in animals and plants organisms are stimuli attracting searchers to work through different ways for identifying new proteins and diagnose their functions for organisms.

Proteins are organic macromolecules made up of linear chains of amino acids joined together by peptide bonds between carboxyl and amino groups of adjacent amino acid residues. The amino acid sequence of a protein is determined by the base pair sequence in the gene which codes for the protein. There are twenty standard amino acids [2].

The total complement of proteins present at a time in a cell is known as its proteome, and the study of such large scale data sets defines the field of proteomics, named by analogy to the related field of genomics. Trough proteomics ways, different proteins of cell or organism can be identified.The key experimental techniques in proteomics include 2-D electrophoresis, which allows the separation of large number of proteins, mass spectrometry, which allows rapid high throughput identification of proteins and sequencing of peptide after in gel digestion, protein microarrays, which allow the detection of the relative levels of a large number of proteins present in a cell etc. These techniques are combined with bioinformatics. The field of bioinformatics seeks to assemble, annotate and analyze genomic and proteomic data applying computational techniques [3].

High resolution 2-D gels can reveal virtually all proteins present in a cell or tissue at any given time, including post-translationally modified proteins. Changes in the expression and structure of most cellular proteins caused by differentiation or external stimuli can be displayed and identified using 2-D protein gels. The proteome